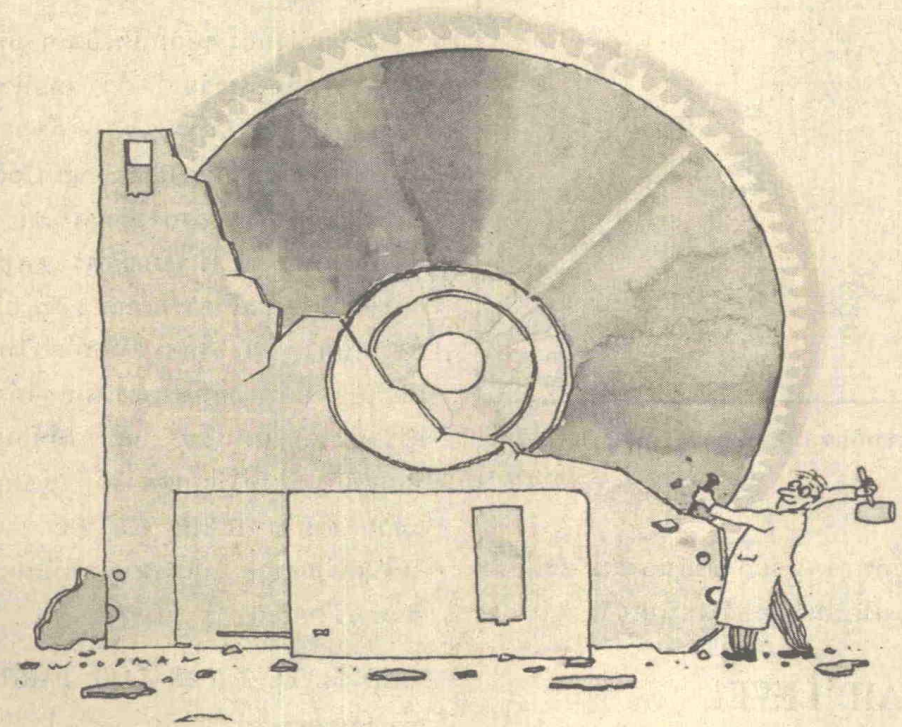
ดิสค์อะเรย์
: ระบบจานบันทึกแบบใหม่



สุดยอดของระบบดิสค์ที่ช่วยเพิ่มประสิทธิภาพในการทำงาน และให้ความเชื่อถือทางด้านข้อมูลสูง

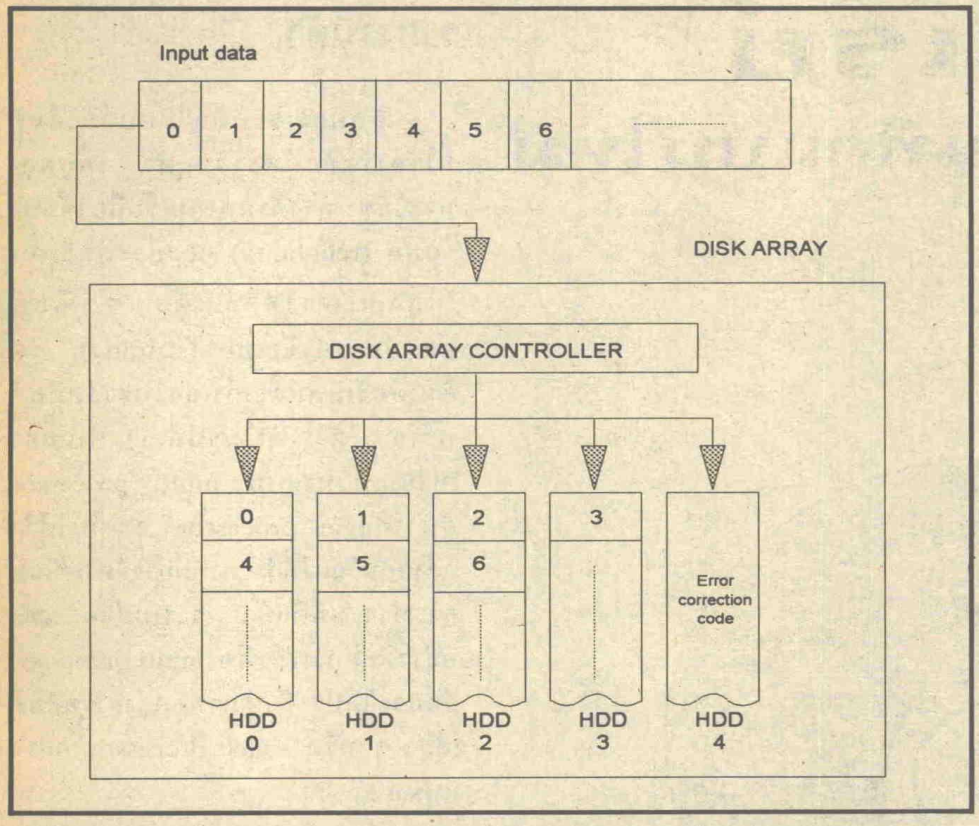
Disk array เป็นเรื่องที่ค่อนข้างใหม่สำหรับวงการคอมพิวเตอร์เมืองไทยและที่จริงแล้วหลักการเกี่ยวกับเรื่องนี้เพิ่งมีมาไม่นานนัก คือประมาณปี 1987 ศาสตราจารย์เดวิด แพทเตอร์สัน และกลุ่มทำงานของเขาที่มหาวิทยาลัย UC.Berkeley ได้บัญญัติศัพท์คำว่า "RAID" หรือคำเต็มว่า Redundancy Arrays of Indexpensive Disks ขึ้นมา ซึ่งเป็นหลักการของดิสค์อะเรย์นี้เอง

รายงานตัว

ดิสค์อะเรย์ก็คือระบบของดิสค์ที่ใช้ฮาร์ดดิสค์หลาย ๆ ตัว เพื่อผลประโยชน์ทางด้านความเชื่อถือได้ของข้อมูล (reliability) และประสิทธิภาพให้สูงกว่าฮาร์ดดิสค์เดี่ยว ๆ ดิสค์อะเรย์จะแบ่งข้อมูล (striping), จัดอินเตอร์เลสและเก็บลงในฮาร์ดดิสค์หลาย ๆ ตัว (ดังในรูปที่ 1) หากลองเปรียบเทียบระบบ multi processor กับ single processor จะเห็นได้ว่าคล้ายกันในแง่ที่ว่า หากมีการจัดระเบียบการทำงานที่ดีแล้ว ระบบช่วยกันทำอย่างดิสค์อะเรย์หรือ multi processor ย่อมจะได้ประสิทธิภาพดีกว่าฮาร์ดดิสค์เดี่ยว ๆ หรือ single processor อย่างแน่นอน

สมมติว่าข้อมูลต่อเนื่องจำนวนมากชุดหนึ่งถูกอ่านหรือเขียนลงในดิสค์อะเรย์อย่างเช่นในรูปที่ 1 ก็จะสามารถกระทำได้พร้อมกันในทุกฮาร์ดดิสค์ และอัตราการส่งผ่านข้อมูลโดยรวมแล้วย่อมจะดีขึ้น ถ้าฮาร์ดดิสค์แต่ละตัวทำงานอิสระกันในการเข้าถึงข้อมูลจำนวนเล็ก ๆ อย่างเช่นอ่านข้อมูลกลุ่ม 5 แล้วก็ 0 (ดูในรูปที่ 1) การเข้าถึงนี้ก็อาจเกิดขึ้นได้พร้อม ๆ กันเลยทั้งกลุ่ม 5 และกลุ่ม 0 ไม่ต้องรอการทำงานเรียงเป็นลำดับ (sequential) เหมือนฮาร์ดดิสค์เดี่ยว ๆ ทั่วไป แน่นอน ประสิทธิภาพการทำงานโดยรวมของดิสค์อะเรย์ย่อมจะดีกว่า

นอกจากนี้เมื่อข้อมูลถูกแบ่งออกมาก็จะมีการเติมรหัสที่เก็บข้อมูลสำหรับตรวจสอบความถูกต้องและแก้ไขด้วย บนฮาร์ดดิสค์หลาย ๆ ตัวเหล่านี้ ใน



รูปที่ 1 การกระจายข้อมูลลงบนดิสค์อะเรย์ และการเก็บข้อมูลเพื่อตรวจสอบและแก้ไข

รูปที่ 1 ฮาร์ดดิสค์หมายเลข 4 จะเก็บข้อมูลเพื่อตรวจสอบและแก้ไขสำหรับข้อมูลในฮาร์ดดิสค์หมายเลข 0 ถึง 3 แม้ว่าฮาร์ดดิสค์ที่เก็บข้อมูลจริง (0-3) ตัวใดตัวหนึ่งจะเสียไป ข้อมูลที่หายไปสามารถกู้กลับมาได้ด้วยข้อมูลที่เก็บไว้ในฮาร์ดดิสค์ตัวอื่น ๆ

แน่นอนที่สุด การเก็บข้อมูลเพื่อตรวจสอบและแก้ไขเหล่านี้จะทำให้ประสิทธิภาพลดลง แต่ละครั้งที่ข้อมูลส่วนหนึ่งถูกเขียนทับ อย่างเช่นข้อมูลส่วนที่ 1 ใน 0-3 ถูกเปลี่ยนแปลง ค่าที่เก็บอยู่ในส่วนตรวจสอบและแก้ไขในฮาร์ดดิสค์ 4 จะต้องถูกคำนวณและเขียนใหม่ด้วย นั่นหมายความว่า ตัวเนื้อข้อมูลเก่าและส่วนตรวจสอบและแก้ไขนี้จะต้องถูกอ่านขึ้นมาทั้ง 2 ส่วน ก่อนที่ข้อมูลใหม่จะถูกเขียนลงไป ซึ่งถือว่าส่วนนี้เป็นโอเวอร์เฮดนั่นเอง

RAID LEVEL

เทคนิคของการทำดิสค์อะเรย์จะแบ่งออกเป็นหลายระดับตามตัวเลขที่ระบุหลัง RAID หลายคนสับสนอยู่ว่าตัวเลขยิ่งมาก หมายถึงการป้องกันรักษาข้อมูลจะยิ่งสูงตามไปด้วย ซึ่งความจริงแล้วไม่ใช่ แพทเตอร์สัน ได้แบ่งการทำดิสค์อะเรย์ออกเป็น 5 ระดับ คือ

RAID LEVEL 1 (DISK MIRRORING)

เป็นการสำรองข้อมูลของดิสค์หนึ่งไปไว้ในอีกตัวหนึ่งแบบตัวต่อตัวเลย จะสามารถป้องกันข้อมูลจากความเสียหายของดิสค์ตัวใดตัวหนึ่งได้ และการอ่านข้อมูลจะดีขึ้น แต่การเขียนข้อมูลจะช้าลง เนื่องจากต้องเขียนลงดิสค์ทั้งคู่ ลักษณะอย่างนี้พบได้ทั้งในเน็ตแวร์หรือแลนแมนเนเจอร์ ในนามของ DISK MIRRORING นั่นเอง

RAID LEVEL 2 (BIT INTERLEAVING WITH HAMMING CODE CORRECTION)

จะแบ่งข้อมูลออกเป็นบิต ๆ เก็บไว้ในฮาร์ดดิสค์หลาย ๆ ตัว โดยดิสค์ชุดหลังจะเก็บบิตเพื่อตรวจแก้ความผิดพลาดที่เรียกว่า Hamming Code Correction ซึ่งต้องการฮาร์ดดิสค์ 3 ตัวเพื่อเก็บโค้ดสำหรับการตรวจสอบและแก้ไขข้อมูลในฮาร์ดดิสค์ 4 ตัว เมื่อข้อมูลบิตไหนหรือในฮาร์ดดิสค์ตัวไหนเสียไป บิตที่เหลือสามารถบอกได้ว่าฮาร์ดดิสค์ตัวไหนเสีย และรวมข้อมูลที่เหลือกันเป็นค่าที่ถูกต้องได้ด้วย ปกติแล้วจะพบกับ RAID LEVEL 2 นี้ได้ในเครื่องพวกเมนเฟรมหรือมินิคอมพิวเตอร์ใหญ่ ๆ

RAID LEVEL 3 (PARALLEL PARITY BIT CHECKING)

วิธีนี้จะคล้าย ๆ กับ LEVEL 2 ที่แบ่งข้อมูลเป็นบิต ๆ ลงฮาร์ดดิสค์แต่ละตัว แต่ LEVEL 3 นี้ต้องการเพียงฮาร์ดดิสค์ตัวเดียวในการเก็บพาริตี้ (parity bit) โดยไม่ว่าข้อมูลจะมีขนาดกี่บิตก็ตาม เมื่อเครื่องคอมพิวเตอร์อ่านหรือเขียนข้อมูล 1 เซกเตอร์ในฮาร์ดดิสค์แบบ RAID 3 ฮาร์ดดิสค์ทุกตัวจะถูกใช้งานพร้อม ๆ กัน (ดังในรูปที่ 2) ซึ่งเป็นการเพิ่มความเร็วในการส่งผ่านข้อมูล เหมาะมากสำหรับการเก็บข้อมูลมาก ๆ เช่น งานอิมเมจโปรเซสซิง

RAID LEVEL 4 (SECTOR STRIPING)

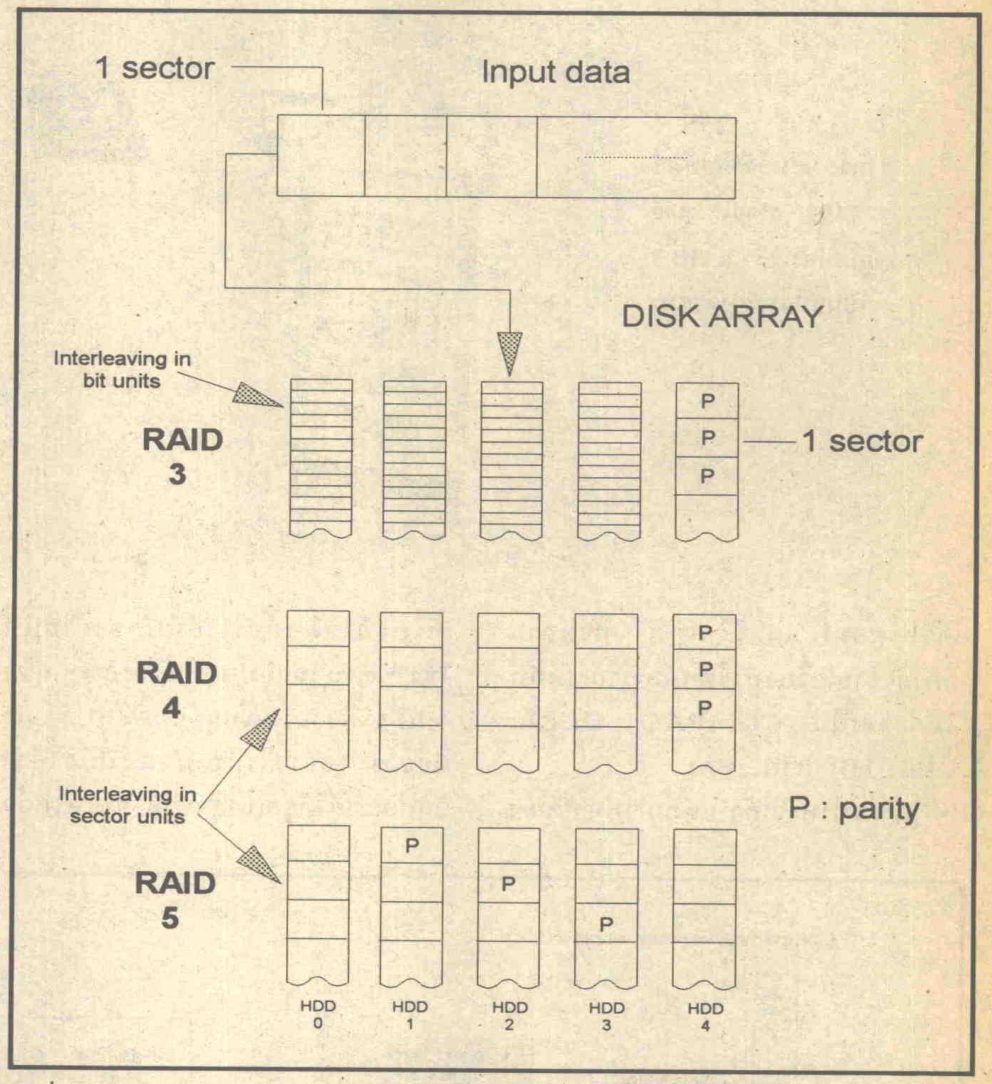
การแบ่งอินเทอร์ลีฟจะถูกควบคุมในระดับเซกเตอร์ของฮาร์ดดิสค์ ซึ่งฮาร์ดดิสค์สามารถใช้งานได้อิสระจากกัน และกระบวนการอ่านการเขียนจะกระทำได้พร้อม ๆ กัน โดยฮาร์ดดิสค์ตัวท้ายสุดจะเก็บพาริตี้เพื่อตรวจสอบความถูกต้องคล้ายกับในระดับ 3 แต่เพราะการเก็บพาริตี้ไว้ที่ฮาร์ดดิสค์ตัวท้ายนี้เอง อาจทำให้เกิดสภาวะคอขวด (bottleneck) ได้ในตอนที่เขียนข้อมูล เนื่องจากจะต้องมีการแก้ไขพาริตี้ และพาริตี้เหล่านี้ก็จะไปรวมกันที่ฮาร์ดดิสค์ตัวเดียวกันหมด ซึ่งจะได้รับการแก้ไขใน RAID LEVEL 5

RAID LEVEL 5 (BLOCK STRIPING WITH DISTRIBUTED PARITY)

หลักการคล้ายกับระดับ 4 ตรงที่จัดการเข้าถึงฮาร์ดดิสค์แบบเซกเตอร์ แต่ดีกว่าในแง่ที่ว่าพาริตี้จะสามารถกระจายอยู่ในฮาร์ดดิสค์ตัวใดก็ได้ ทำให้การเขียนข้อมูลหลาย ๆ ชุดซึ่งต้องมีการอ่านและแก้ไขพาริตี้สามารถกระทำพร้อม ๆ กันได้

นอกจากนี้แล้ว การแบ่งเป็นกลุ่ม (striping) ที่ไม่มีการตรวจสอบและแก้ไขข้อมูลก็อาจเรียกว่า RAID 0 ด้วย หรือถ้าทำ Mirrored disk พร้อมกับการแบ่งกลุ่ม ก็จะได้ทั้งความน่าเชื่อถือและประสิทธิภาพ ซึ่งเรียกวิธีการนี้ว่า RAID 10

สรุปในรูปที่ 2 บอกความแตกต่างระหว่าง RAID แต่ละระดับใน RAID 2 และ RAID 3 ข้อมูลจะแบ่งเป็นบิต ส่วน RAID 4 และ RAID 5 แบ่งเป็นเซกเตอร์ของฮาร์ดดิสค์ (โดยทั่วไปขนาดหน่วยละ 512 ไบต์) RAID 2 ใช้วิธี Hamming Error Correction Code ในขณะที่ RAID 3, 4 และ 5 ใช้พาริตี้ และ RAID 2, 3, 4 ใช้ฮาร์ดดิสค์ชุดหลังเป็นที่เก็บตัวตรวจสอบแก้ไข ส่วน RAID 5 จะสามารถกระจายไปอยู่ที่ฮาร์ดดิสค์ตัวไหนก็ได้

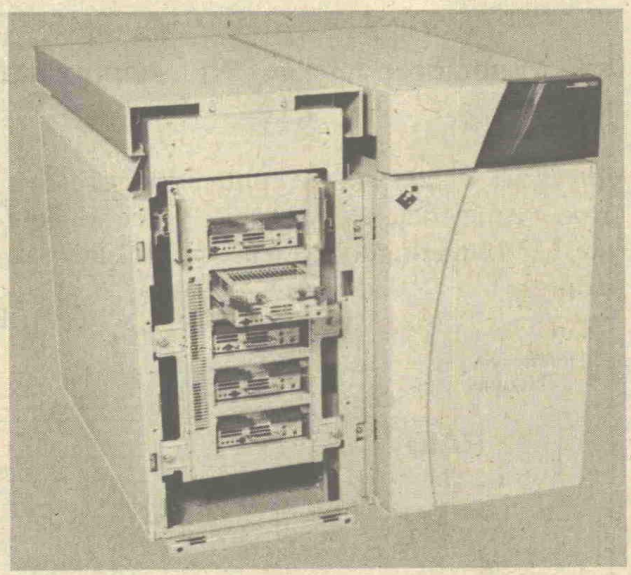


รูปที่ 2 เปรียบเทียบข้อแตกต่างระหว่างดิสค์อะเรย์ RAID ต่าง ๆ กัน

ใช้เป็นศูนย์บริการข้อมูล

ในปัจจุบันดิสค์อะเรย์ไม่ได้อยู่เพียงในรูปแบบของทฤษฎีแล้ว เพราะได้มีการนำดิสค์อะเรย์มาใช้ในวงการค้าจริง ๆ เครื่องเอ็นอีซีตระกูล System 3100 ที่เริ่มจำหน่ายเมื่อประมาณปี 1992 ได้ใช้ดิสค์อะเรย์สำหรับรุ่นใหญ่ทั้ง 9 รุ่น และจะใช้ดิสค์อะเรย์อีกในระบบปฏิบัติการยูนิกซ์ในอนาคตอันใกล้นี้อีกด้วย

นอกจากนี้บริษัทโซนี่ก็เริ่มจำหน่ายเซอร์ฟเวอร์ยูนิกซ์ที่ใช้ดิสค์อะเรย์เป็นอุปกรณ์มาตรฐานด้วย (ดังในรูปที่ 3) สำหรับในอเมริกาแล้ว บริษัทที่จำหน่าย



รูปที่ 3
เครื่องเอ็นอีซีซิสเต็มส์
3100 Model A65
ที่มีดิสค์อะเรย์ RAID 3
เป็นอุปกรณ์มาตรฐาน

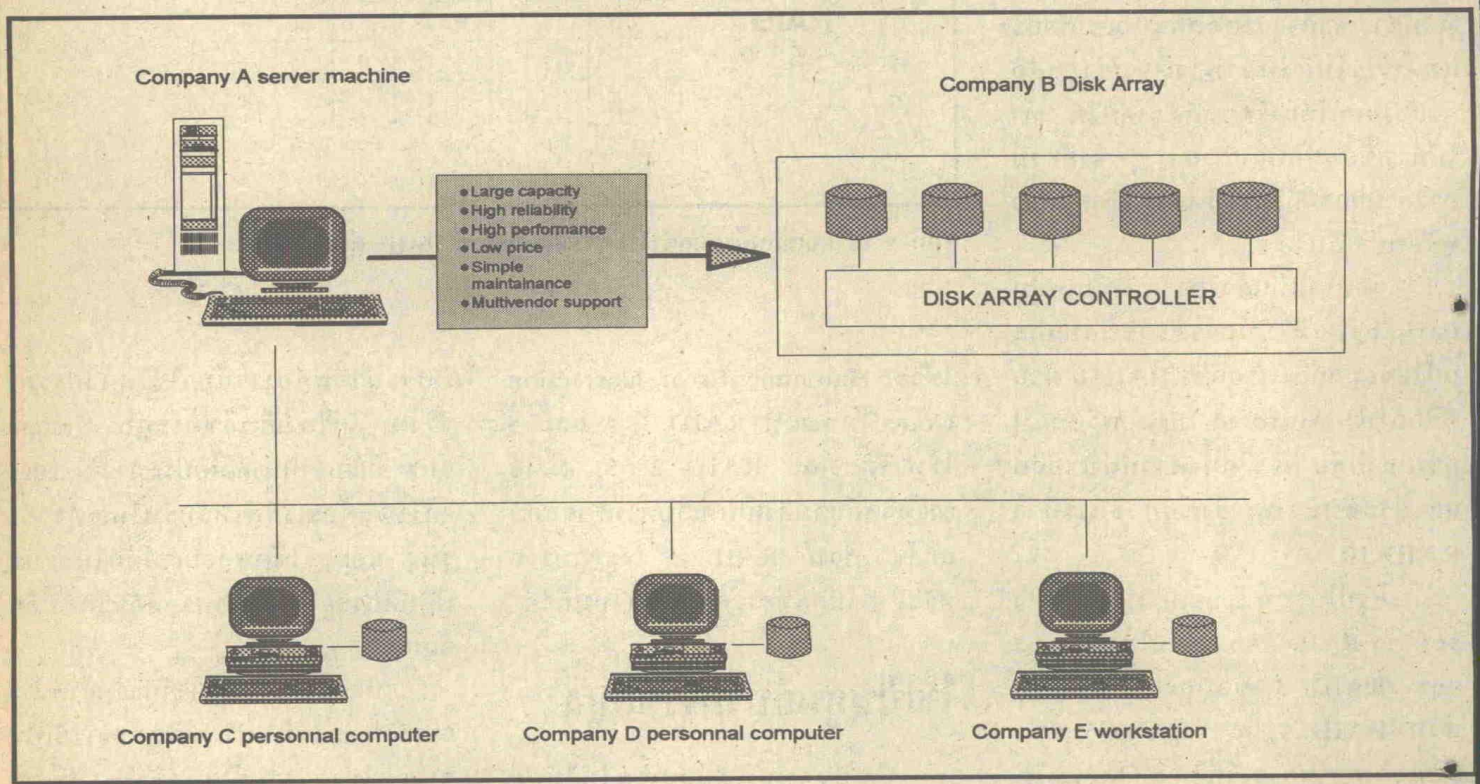
เครื่องคอมพิวเตอร์ใหญ่ ๆ ทั้งหลายก็กำลังขายเครื่องที่ใช้ดิสค์อะเรย์เช่นกัน ไม่ว่าจะเป็น COMPAQ, DELL, DEC, HP หรือ IBM

ส่วนใหญ่แล้ว กลุ่มเป้าหมายของเครื่องที่ใช้ดิสค์อะเรย์นี้มักจะเป็นผู้ที่ใช้งานด้านวิจัยอิมเมจโปรเซสซิง หรืองานที่จะต้องอ่านเขียนข้อมูลมาก ๆ และบ่อย ๆ อย่างไรก็ตามขณะนี้ถือได้ว่าดิสค์อะเรย์ได้รุกเข้าสู่งานด้านธุรกิจแล้ว

และยุคนี้ก็ถือได้ว่าเป็นยุคของระบบ DISTRIBUTED COMPUTER ซึ่งเชื่อมโยงคอมพิวเตอร์ทั้งหลายเข้ากันเป็นเครือข่าย ตัวที่จะเป็นศูนย์บริการข้อมูลหรือเครื่องที่เป็นเซอร์ฟเวอร์ (server) จำเป็นต้องมีระบบไฟล์ขนาดใหญ่ เชื่อถือได้ ประสิทธิภาพสูง ราคาถูกและหาง่าย ดูแลรักษาง่าย คุณสมบัติทั้งหลายเหล่านี้จึงทำให้ดิสค์อะเรย์กำลังเป็นที่สนใจอยู่ในขณะนี้ (ดูรูปที่ 4)

ข้อดีของดิสค์อะเรย์

ข้อดีที่เห็นได้ชัดคือความน่าเชื่อถือของข้อมูลเมื่อเปรียบเทียบกับฮาร์ดดิสค์เดี่ยว ๆ แล้วต่างกันมาก ถ้าเครื่องไม่ได้ถูกไฟไหม้หรือภัยพิบัติอย่างอื่นแล้ว โอกาสที่ข้อมูลจะสูญหายมีเพียงหนึ่งครั้งในช่วงเวลานับหมื่นปีทีเดียว



รูปที่ 4 ดิสค์อะเรย์ในสภาพแวดล้อมของ distributed computing

ต้นฉบับไม่มีหน้า 207-208

No This Page 207-208 In Original

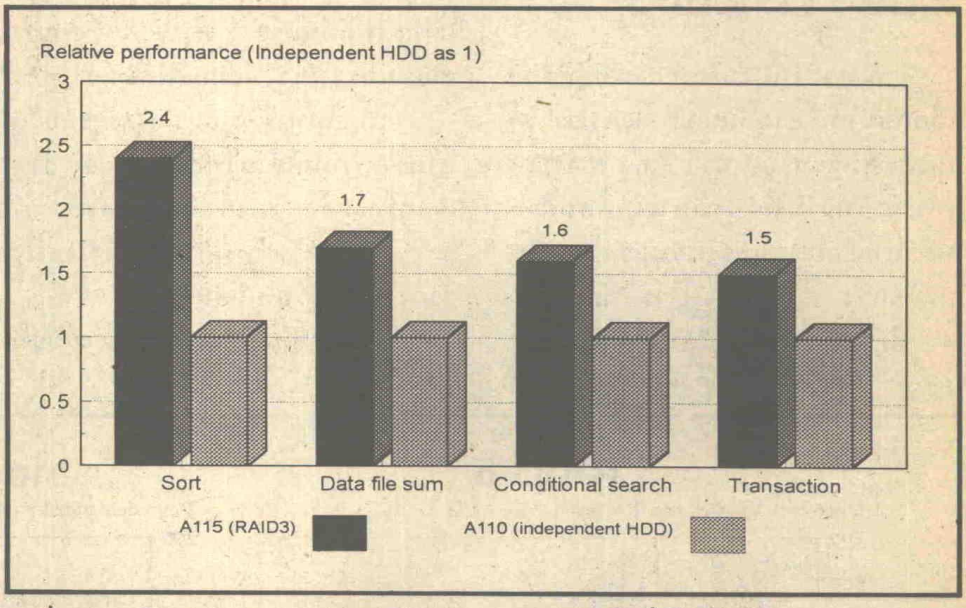
ตารางที่ 1 ตารางแสดงคุณสมบัติของดิสค์อะเรย์ในทางทฤษฎีกับที่เป็นจริง ซึ่งยังพบปัญหาอีกมาก

	Theory	Reality
Capacity	สามารถสร้างให้มีความจุมากกว่าฮาร์ดดิสค์เดี่ยว ๆ ได้หลายเท่า	ความจุของฮาร์ดดิสค์เดี่ยว ๆ เพิ่มขึ้นอย่างรวดเร็วและเพียงพอ
Reliability	หากไม่มีอุบัติเหตุอย่างเช่นไฟไหม้ โอกาสข้อมูลเสียหายมีเพียงหนึ่งในหมื่นปี	ยังมีใช้ไม่มากนัก ข้อสรุปจึงยังไม่แน่นอน และมีปัจจัยอื่นมีมากนอกเหนือจากตัวดิสค์อะเรย์เอง
Performance	การใช้งานกับแอพพลิเคชันอย่างเหมาะสม จะได้ประสิทธิภาพสูงกว่าฮาร์ดดิสค์เดี่ยว ๆ	ระบบคอมพิวเตอร์ส่วนใหญ่ยังไม่มีฟังก์ชันการทำงานที่จะใช้ประโยชน์ดิสค์อะเรย์ได้ สูงสุด และยังไม่อาจระบุได้ว่าแอพพลิเคชันใดเหมาะกับ RAID LEVEL ใด
Price	สามารถสร้างได้ราคาถูกกว่า mirrored disks	mirrored disk จะถูกกว่าที่ขนาดความจุต่ำ
Multivendor support	สามารถใช้งานได้กับอินเตอร์เฟสมาตรฐานต่าง ๆ ได้	ยังต้องการระบบดีไวซ์ไดรเวอร์อีกมาก

เรื่องของราคาก็สำคัญ ปัจจุบันเซอร์ฟเวอร์ที่ว่าไว้ใจได้ก็มักจะใช้ระบบ MIRRORED DISK ซึ่งจะใช้ฮาร์ดดิสค์อีกหนึ่งเท่าตัวในการทำสำรองข้อมูล เมื่อเทียบกับดิสค์อะเรย์ซึ่งต้องการฮาร์ดดิสค์ที่เป็นตัวตรวจสอบน้อยกว่ามาก

ข้อดีอีกอย่างก็คือ ในทางทฤษฎีแล้วดิสค์อะเรย์จะมีความเร็วในการอ่านการเขียนข้อมูลมากกว่าฮาร์ดดิสค์เดี่ยว ๆ มาก เนื่องจากสามารถเข้าถึงฮาร์ดดิสค์แต่ละตัวได้พร้อม ๆ กัน

หากแต่ในทางปฏิบัติแล้ว ขณะนี้พบว่าคุณสมบัติของดิสค์อะเรย์ยังไม่เป็นไปตามทฤษฎีนัก (ดูในตาราง) สินค้าที่ออกมายังตลาดยังมีประสิทธิภาพและราคาไม่ดีกว่าเท่าใดนัก อนาคตของดิสค์อะเรย์จะไปรอดหรือไม่นั้น ขึ้นอยู่กับว่าผู้ผลิตสามารถแก้ไขข้อบกพร่องต่าง ๆ ที่ยังเหลืออยู่ในปัจจุบันได้หรือไม่ เช่น ในเรื่องของราคา ปัจจุบันพบว่าที่ความจุไม่เกิน 4 จิกะไบต์ (4000 ล้าน) ราคาของการทำ MIRRORED DISK จะถูกกว่าการใช้ดิสค์อะเรย์ ทั้งนี้ก็เนื่องจากตัวคอนโทรลเลอร์ และปริมาณการผลิตที่ยังไม่มากนัก



รูปที่ 5 ผลการเปรียบเทียบระหว่างเอ็นอีซี A110 (ฮาร์ดดิสค์เดี่ยว ๆ)กับ A115 (RAID 3)

อีกอย่างหนึ่ง เรื่องของประสิทธิภาพปัจจุบันยังไม่อาจกล่าวได้ว่าดิสค์อะเรย์มีประสิทธิภาพเหนือกว่าฮาร์ดดิสค์เดี่ยว ๆ เพราะว่าต้องขึ้นอยู่กับโครงสร้างของดิสค์อะเรย์เอง และผู้ใช้งานมันโดยตรงก็คือตัวระบบปฏิบัติการของเซอร์ฟเวอร์ที่จะต้องรู้จักและใช้งานมันเป็น และเหมาะกับงานประยุกต์หรือแอพพลิเคชันข้างบน และในปัจจุบัน ก็มีพีซีหรือเวอร์กสเตชันจำนวนน้อยมาก ที่รู้จักและสามารถใช้งานดิสค์อะเรย์ระดับ RAID 5 ได้เต็มประสิทธิภาพ

หรืออย่างอินเตอร์เฟสพวก SCSI-2 ดิสค์อะเรย์ก็พบปัญหาอีก เนื่องจากธรรมชาติของดิสค์อะเรย์ต้องการการทำงานแบบขนาน (parrallel) ต่อตัวฮาร์ดดิสค์ เพื่อผลทางด้านความเร็วของการอ่านและการเขียน แต่ SCSI-2 ไม่สามารถสั่งงานแบบขนานหรือพร้อมกันได้ ดังนั้นคุณสมบัติระดับ RAID 5 จึงเกิดอุปสรรค สำหรับ TAGGED COMMAND QUEUING ที่ SCSI-2 มีให้ก็เป็นวิธีการแก้ปัญหาอย่างหนึ่ง แต่ในปัจจุบันการทำงานลักษณะอย่างนี้พบได้น้อยมากในพีซีหรือเวอร์กสเตชัน

RAID 5 v.s. RAID 3

การจะให้ประสิทธิภาพที่ดีก็ต้องเลือกชนิดหรือระดับของ RAID ให้เหมาะสมด้วย โดยทั่วไปแล้ว RAID 3 เหมาะกับงานที่ต้องมีการเคลื่อนย้ายหรือถ่ายโอนข้อมูลจำนวนมาก ๆ ในแต่ละครั้ง ส่วน RAID 5 เหมาะกับงานที่อ่านเขียนข้อมูลไม่มากนักแต่บ่อยครั้ง ฟังดูคล้าย ๆ กับว่า RAID 3 เหมาะกับงานด้านการคำนวณทางวิทยาศาสตร์ และอิมเมจโปรเซสซิงและ RAID 5 เหมาะกับงานทางด้านธุรกิจ

อย่างไรก็ตาม RAID 3 ก็สามารถใช้กับงานด้านธุรกิจอย่างมีประสิทธิภาพเช่นกัน อย่างเช่น ดึงข้อมูลจากระบบฐานข้อมูลขนาดโต ๆ ซึ่งก็เป็นจุดแข็งของ RAID 3 (ดังในรูปที่ 5) การทดสอบในที่นี้มีการทำงาน 4 ลักษณะ คือ

1) อ่านไฟล์ขนาด 1 ล้านเรคอร์ด เรียงจากมากไปน้อยแล้วเขียนลงอีกไฟล์หนึ่ง

2) หาผลรวมจากไฟล์ 2 ไฟล์ที่แต่ละไฟล์มีขนาด 1 ล้านเรคอร์ด แล้วเขียนผลรวมลงไฟล์อีกไฟล์หนึ่ง

3) ค้นหาข้อมูล 1 แสนเรคอร์ดตามข้อกำหนดจากไฟล์ขนาด 2 ล้านเรคอร์ด แล้วแสดงผลออกทางจอ

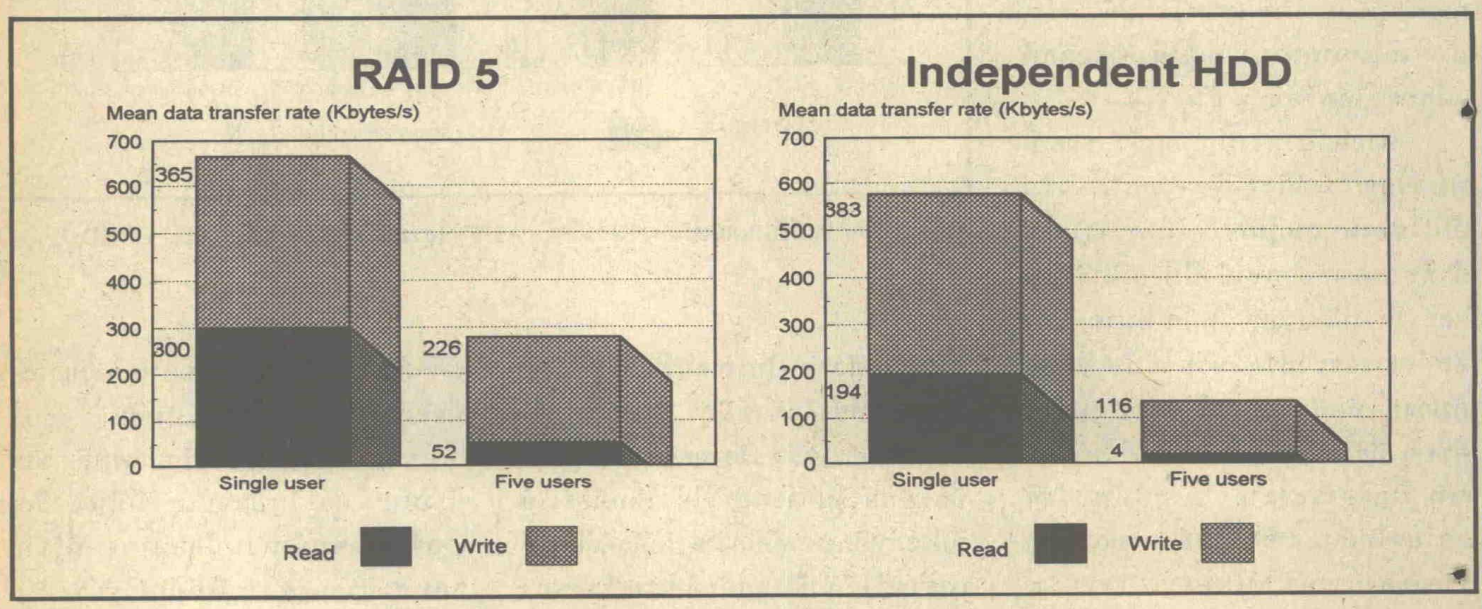
4) แก้ไขข้อมูลในไฟล์อันหนึ่งจาก 80 เวอร์กสเตชัน

ในการที่ดิสค์อะเรย์จะมีการเขียนข้อมูลลงไป จะต้องเสียเวลาในการอ่านข้อมูลเก่าและพาริตี้ขึ้นมาก่อน ใน RAID 3 สามารถลดโอเวอร์เฮดส่วนนี้ลงได้ด้วยการกำหนดขนาดของข้อมูลที่จะเข้าถึงแต่ละครั้งให้มีขนาดที่เหมาะสม และถ้าการหมุนของดิสค์และการเคลื่อนหัวอ่านของฮาร์ดดิสค์ที่เป็นส่วนประกอบสอดคล้องกัน ประสิทธิภาพโดยรวมของดิสค์อะเรย์จะเหนือกว่าฮาร์ดดิสค์เดี่ยว ๆ อย่างแน่นอน

สำหรับ RAID 5 ซึ่งเหมาะกับงานที่มีการอ่านเขียนบ่อย ๆ จึงเหมาะกับงานในระบบเครือข่าย ซึ่งทำงานในลักษณะแบบผู้ใช้หลายคน (MULTI-USER) เนื่องจากฮาร์ดดิสค์แต่ละตัวสามารถถูกอ่านและเขียนเป็นอิสระจากกัน (แต่สอดคล้องกัน) การเขียนไฟล์หลาย ๆ ไฟล์พร้อมกันจึงดูไม่เป็นเรื่องลำบากนัก

ใช้ได้ดีกับระบบเครือข่าย

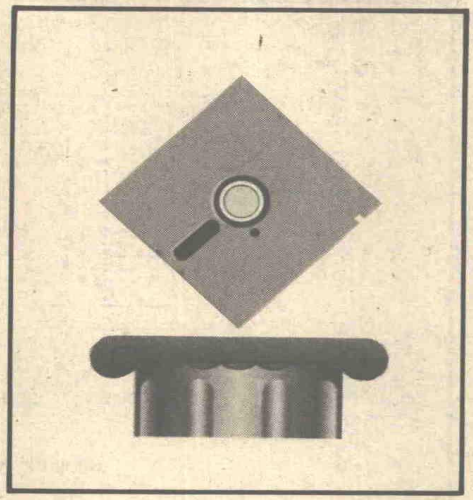
ในสภาวะแวดล้อมของระบบเครือ



รูปที่ 6 ผลการทดสอบ RAID 5 ในระบบเครือข่าย ระหว่างผู้รับบริการ 1 คนกับ 5 คน

ข่ายซึ่งมีผู้ใช้ (CLIENT) จำนวนหนึ่งขอใช้บริการข้อมูลจากเซอร์ฟเวอร์ จะมีคำสั่งจากผู้ใช้หรือผู้ขอรับบริการให้อ่านหรือเขียนไฟล์ของเซอร์ฟเวอร์ โดยเป็นไปในลักษณะสุ่ม (RANDOM) และความอิสระหรือความสุ่มนี้จะเพิ่มขึ้นตามจำนวนของผู้ใช้

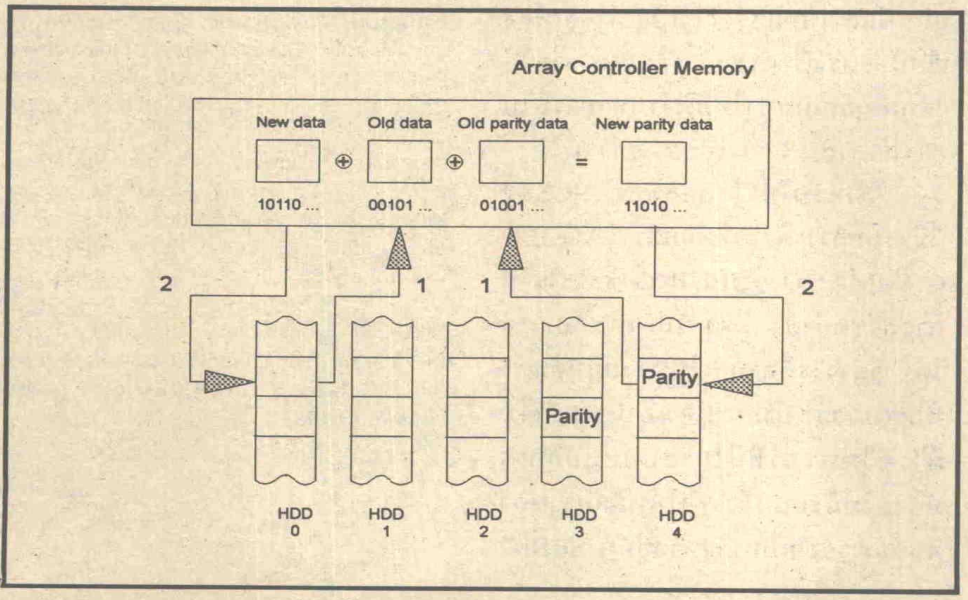
จากการสำรวจจาก NIPPON EUROTEC ถึงประสิทธิภาพของ RAID 5 ในระบบเครือข่าย โดยใช้เซอร์ฟเวอร์ขนาด 33 เมกะเฮิรตซ์ i486 หน่วยความจำขนาด 33 เมกะไบต์ กับผู้ใช้ (CLIENT) เป็นเครื่องตั้งแต่ 80386-20 เมกะเฮิรตซ์ จนถึง i486DX2-50 เมกะเฮิรตซ์ โดยใช้ระบบปฏิบัติการเครือข่ายเน็ตแวร์ 3.11 บนอีเธอร์เนต ไฟล์ขนาด 65 เมกะไบต์ หลาย ๆ ไฟล์ ถูกส่งผ่านระหว่างผู้ใช้กับเซอร์ฟเวอร์ เพื่อเปรียบเทียบดิสค์อะเรย์ขนาด 1.2 จิกะไบต์ ของ MYLEX CORP ของอเมริกา กับฮาร์ดดิสค์เดี่ยว ๆ ขนาด 1.3 จิกะไบต์ ของ MAXTOR CORP ของอเมริกา ผลออกมาตามรูปที่ 6 ซึ่ง



เห็นได้ว่าประสิทธิภาพของดิสค์อะเรย์จะดีกว่าฮาร์ดดิสค์เดี่ยว ๆ สำหรับเครือข่ายที่มีผู้ใช้ 5 คนมากกว่าเครือข่ายที่มีผู้ใช้เพียงคนเดียว

การเพิ่มประสิทธิภาพ

ผู้ผลิตเครื่องที่ใช้ดิสค์อะเรย์พบว่าปัญหาใหญ่ของดิสค์อะเรย์ก็คือ ความช้าที่เกิดตอนเขียนข้อมูล ซึ่งจะต้องมีการคำนวณพาริตี้ใหม่นั่นเอง (ดังในรูปที่ 7) หนทางแก้ไขก็ทำได้โดยต้องหาทางลดเวลาที่ใช้ในการเขียนนี้ และพยายามเขียนข้อมูลเป็นชุดใหญ่ ๆ ครั้งละมาก ๆ



รูปที่ 7 ขั้นตอนการอ่านพาริตี้เก่าขึ้นมาคำนวณใหม่แล้วเขียนทับลงไป

ใช้บัฟเฟอร์แคช

ผู้ผลิตดิสค์อะเรย์ RAID 5 เห็นว่าการที่จะแก้ปัญหาโอเวอร์เฮดในการเขียนข้อมูลจะต้องใช้หลักการของบัฟเฟอร์แคช (BUFFER CACHE) ข้อมูลจากเซอร์ฟเวอร์ที่จะเขียนจะถูกส่งลงบัฟเฟอร์แคชนี้ และดิสค์อะเรย์ก็จะบอกไปยังเซอร์ฟเวอร์ว่าการเขียนนั้นเรียบร้อยแล้ว ทำให้เซอร์ฟเวอร์สามารถทำงานอย่างอื่นต่อไปได้เลย ถ้ามีการแก้ไขข้อมูลที่ยังอยู่ในบัฟเฟอร์แคช ข้อมูลนั้นก็จะแก้ไขในหน่วยความจำหรือบัฟเฟอร์แคชนั่นเอง ทำให้ประสิทธิภาพดีขึ้น การเขียนลงดิสค์ก็จะใช้เวลาน้อยลง

ขนาดลอจิกคอลบล็อก

ส่วนประสิทธิภาพของ RAID 3 ก็ติดปัญหาที่โอเวอร์เฮดในการเขียนเช่นกัน ซึ่งจะขึ้นอยู่กับขนาดของข้อมูลเป็นลอจิกคอลบล็อก (LOGICAL BLOCK) ที่ระบบปฏิบัติการจัดการ ขนาดของบล็อกนี้จะต้องเป็นจำนวนเท่าของขนาดบล็อกสำหรับดิสค์อะเรย์หรือที่เรียกว่าบล็อกทางกายภาพ (PHYSICAL BLOCK)

ในรูปที่ 8 เป็นตัวอย่างสมมติระบบที่ใช้ RAID 3 ระบบหนึ่งประกอบด้วยฮาร์ดดิสค์ 5 ตัว ขนาดข้อมูลที่ใช้ในการเข้าถึงฮาร์ดดิสค์แต่ละตัวเป็นบล็อกหรือความยาวเซกเตอร์ขนาด 512 ไบต์ RAID 3 จะอ่านหรือเขียนฮาร์ดดิสค์

ทั้ง 5 ตัวนี้พร้อม ๆ กัน เนื่องจากจะมีฮาร์ดดิสค์ที่เป็นพาริตี้ 1 ตัว ขนาดของข้อมูลที่ใช้ในการอ่านและการเขียนดิสค์อะเรย์นี้จะเป็น 4 คูณ 512 ไบต์ เท่ากับ 2 กิโลไบต์นั่นเอง

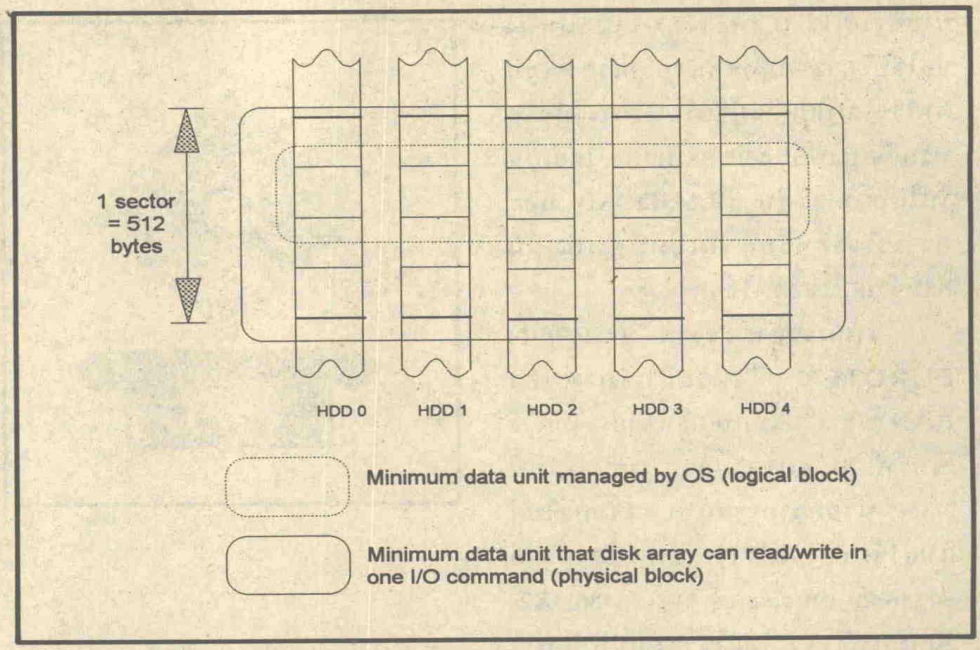
ในการอ่านหรือการเขียนข้อมูลในดิสค์อะเรย์ระบบนี้จึงควรจะใช้ขนาดลอจิกคอลบล็อกที่เป็นจำนวนเท่าของ 2 กิโลไบต์ ในการเขียนข้อมูลบล็อกทางกายภาพที่เป็นที่อยู่ของลอจิกคอลบล็อกที่ต้องการจะถูกแก้ไขและก็เขียนกลับลงไปในดิสค์ บัฟเฟอร์แคชสามารถช่วยเป็นตัวกลางประสานความแตกต่างของบล็อกทางกายภาพและลอจิกคอลบล็อกได้ และช่วยลดโอเวอร์เฮดในการเขียนลงในที่สุด

ชุดของคำสั่ง

การปรับปรุงวิธีการสั่งชุดคำสั่งในการอ่านหรือการเขียนข้อมูล RAID 5 ก็จะช่วยเพิ่มประสิทธิภาพได้มาก โดยคำสั่งอินพุต-เอาต์พุต (I/O) อันถัดไปสามารถสั่งไปได้เลยโดยไม่ต้องรอผลของคำสั่ง I/O อันก่อนหน้า ระบบการเชื่อมต่อหรือการอินเตอร์เฟสระหว่างเซอร์ฟเวอร์กับดิสค์อะเรย์จะต้องสามารถจัดการการส่งคำสั่ง I/O อย่างต่อเนื่องได้ ดิสค์อะเรย์ RAID 5 มีความสามารถทำงานสนองคำสั่ง I/O ในลักษณะแบบขนาน (PARALLEL) แต่ละคำสั่ง I/O จะทำให้เกิดการทำงานต่อฮาร์ดดิสค์อย่างมาก 2 ตัว (ข้อมูลกับพาริตี้) จึงสามารถเพิ่มประสิทธิภาพการทำงานโดยรวมได้ด้วยการทำงานสนองคำสั่ง I/O อื่น ๆ ไปพร้อมกันได้ แทนที่จะปล่อยให้ฮาร์ดดิสค์อื่นว่างงานเฉย ๆ (IDLE)

สำหรับระบบปฏิบัติการหรือ OPERATING SYSTEM แล้ว ที่นิยมใช้ในลักษณะเซอร์ฟเวอร์อย่างนี้คือ ยูนิกซ์และเน็ตแวร์ จะมีปัญหาอยู่เพียงเล็กน้อยตรงเรื่องของการอินเฟส หากไม่ได้ถูกออกแบบมาให้มีความสามารถในการส่งคำสั่ง I/O เป็นชุดอย่างนี้ได้

สำหรับโปรโตคอลของ SCSI-2 ก็มีระบบการทำงานที่เรียกว่า TAGGED COMMAND QUEUING ซึ่งจะใช้ในการส่งชุดคำสั่ง I/O ไปยังดิสค์อะเรย์ได้ อย่างไรก็ตาม ในทางปฏิบัติแล้วมักพบว่าระบบควบคุมอุปกรณ์หรือดีไวซ์ไดรเวอร์ที่มีในระบบปฏิบัติการมักจะไม่มีระบบนี้ ทำให้หลักการของดิสค์อะเรย์ยังเป็นปัญหาอยู่กับ SCSI-2



รูปที่ 8 ตัวอย่างของ RAID 3 ที่มีความยาวเซกเตอร์ 512 ไบต์ และขนาดลอจิกคอลบล็อก 512 ไบต์ในการแก้ไขข้อมูลในกรอบเส้นประจะต้องอ่านข้อมูลทั้งหมดในบริเวณเส้นทึบขึ้นมา

หนังสืออ้างอิง

Patterson, D A, Gibson, G and Kagz, R H, "A Case ror Redundant Arrays of Inexpensive Disks (RAID)," Report no UCB/CSD87/391, Computer Science Div, UC.Berkeley, 1987.

Katz, R H, Gibson, G A and Patterson, D A, "Disk System Architechtures for High Performance Computing," Proceedings of the IEEE, vol 77, no 12, pp 1842-1858, December 1989.

Nikkei Electronics Asia, July 1993, "WS, PC Makers Eye Enhanced disk Arrays for Server Storage," Nikkei Business Publications, July 1993.